

## WE CLAIM:

1. A method of speech recognition, said method comprising the steps of:

receiving audio signals from a speech source;  
receiving video signals from the speech source;  
processing the audio signals and the video signals;  
converting the audio signals and the video signals into recognizable information;  
implementing a task based on the recognizable information.

2. The method of claim 1, wherein the step of receiving the video signals comprises the step of:

receiving video images of lip movements that coincide with the audio signals.

3. The method of claim 1, wherein the step of processing comprises the step of:

processing the audio signals and the video signals in parallel, wherein the video signals coincide with the audio signals.

4. The method of claim 1, further comprising the steps of:

storing the audio signals and the video signals; and  
sending the audio signals and the video signals to a destination source.

5. A speech recognition device, said device comprising:

an audio signal receiver configured to receive audio signals from a speech source;

a video signal receiver configured to receive video signals from the speech source;

a processing unit configured to process the audio signals and the video signals;

a conversion unit configured to convert the audio signals and the video signals to recognizable information;

an implementation unit configured to implement a task based on the

recognizable information.

6. The device of claim 5, wherein the video signal receiver is configured to receive video images of lip movements that coincide with the audio signals.

7. The device of claim 5, wherein the processing unit is configured to process the audio signals and the video signals in parallel, wherein the video signals coincide with the audio signals.

8. The device of claim 5, further comprises:  
a storage unit for storing the audio signals and the video signals; and  
a transmitter for sending the audio signals and the video signals to a destination source.

9. A system for speech recognition, said system comprising:  
a first receiving means for receiving audio signals from a speech source;  
a second receiving means for receiving video signals from the speech source;  
a processing means for processing the audio signals and the video signals;  
a converting means for converting the audio signals and the video signals to recognizable information;  
an implementing means for implementing a task based on the recognizable information.

10. The system of claim 9, wherein the second receiving means receives video images of lip movements that coincide with the audio signals.

11. The system of claim 9, wherein the processing means processes the audio signals and the video signals in parallel, wherein the video signals coincide with the audio signals.

12. The system of claim 9, further comprises:  
a storage means for storing the audio signals and the video signals; and  
a transmission means for sending the audio signals and the video

signals to a destination source.

13. A method of speech recognition, said method comprising the steps of:

receiving audio signals from a speech source;

receiving video signals from the speech source;

processing the audio signals;

converting the audio signals into recognizable information;

processing the video signals when a segment of the audio signals can not be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information;

converting the processed video signals into the recognizable information; and

implementing a task based on the recognizable information.

14. A speech recognition device, said device comprising:

an audio signal receiver configured to receive audio signals from a speech source;

a video signal receiver configured to receive video signals from the speech source;

a first processing unit configured to process the audio signals;

a first conversion unit configured to convert the audio signals to recognizable information;

a second processing unit configured to process the video signals when the audio signals cannot be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information;

a second conversion unit configured to convert the processed video signals into the recognizable information; and

an implementation unit configured to implement a task based on the recognizable information.

15. A system for speech recognition, said system comprising:
- a first receiving means for receiving audio signals from a speech source;
  - a second receiving means for receiving video signals from the speech source;
  - a first processing means for processing the audio signals;
  - a first converting means for converting the audio signals into recognizable information;
  - a second processing means for processing the video signals when a segment of the audio signals can not be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information;
  - a second converting means for converting the processed video signals into the recognizable information; and
  - an implementing means for implementing a task based on the recognizable information.